Computer Engineering and Applications Vol. 8, No. 3, October 2019

Measuring Inter-VM Performance Interference in IaaS Cloud

Kenga Mosoti Derdus[1,*], Vincent Oteke Omwenga[1], Patrick Job Ogao[2]

[1]*Faculty of Information Technology, Strathmore University, Nairobi, Kenya*
[2]*Faculty of Engineering Science and Technology, Technical University of Kenya, Nairobi, Kenya*
**derduskenga@gmail.com*

ABSTRACT

Virtualization has enabled cloud computing to deliver computing capabilities using limited computer hardware. Server virtualization provides capabilities to run multiple virtual machines (VMs) independently in a shared host leading to efficient utilization of server resources. Unfortunately, VMs experience interference from each other as a result of sharing common hardware. The performance interference arises from VMs having to compete for the hypervisor capacity and as a result of resource contention, which happens when resource demands exceed the allocated resources. From this viewpoint, any VM allocation policy needs to take into account VM performance interference before VM placement. Therefore, understanding how to measure performance interference is crucial. In this paper, we propose a simple experimental approach that can be used to measure performance interference in Infrastructure as a Service (IaaS) cloud during VM consolidation.

Keywords: Cloud Computing, Virtual Machines, Performance Interference, Virtualization

1. INTRODUCTION

Nowadays, cloud computing is common in delivering IT service because of its success in delivering the services on pay-as-you-go model. Sharing of resources in cloud computing is made possible by virtualization, where one hardware (host) is broken down into small independent machines (VM), which execute applications independently [1]. Although server virtualization has well-known advantages such as security isolation of VM and efficient use of server resources, it suffers from performance interference as demonstrated in [2] and [3] because it lacks performance isolation mechanisms. In fact, it has been shown that I/O bandwidth can by almost 50% due to inter-VM interference. Therefore, successful VM consolidation is that, which strives to minimize performance interference. To do so, there must be a way of measuring performance interference.

In this paper, we present a simple experimental method that can be used to measure performance inference at different levels of consolidation in CPU-intensive workloads. Although methods of minimizing performance interference have been proposed such as in [4], [5] and [6] , there is no attempt made to show it can be measured explicitly. The remainder of this paper is organized as follows: in section II, we present related works. Section III describes our experimental setup. In section IV, our experimental results and discussion are presented and the paper is concluded in section V.

2. RELATED WORKS

In [4], the authors have proposed a framework, which measures performance interference for CPU and network I/O intensive workloads when they run in parallel. The aim of the experiment is to compare interference between network I/O intensive workloads and CPU intensive workloads when they run in parallel and independently. The authors have concluded that performance interference is lower when the workloads run in parallel as opposed to when they run independently. This deduction is important because it is a factor that should be considered while scheduling application workloads.

In [5], the authors have proposed a system called Cloudscope for predicting interference between co-located workloads in a multi-tenant environment. The authors argue that their application is lightweight as compared to existing approaches, which uses online training causing them to be computationally intensive. Cloudscope predicts workloads interference and then reassigns VM to physical servers, which will cause the least interference. Cloudscope has been implemented in Xen hypervisor. In [6], the authors have proposed a machine learning system for detecting if a web service running in VMs is suffering from performance interference. This approach relies on known historic performance of a workload in given hardware with current performance. Because of the fact that the system uses a machine learning approach and that it is implemented within a VM, it may be computationally intensive.

3. EXPERIMENTAL SETUP

For inter-VM performance interference to be measured, an application workload needs to be executed in a VM hosted in a physical server together with other VMs. In this experiment, we have used 7zip and aio-stress, which are application benchmarks within Phoronix Test Suite (PTS) as the source of workload. PTS is the most compressive open-source benchmarking tool [7], [8]. PTS has a collection test profiles with different but known behaviours. 7zip is a file compression application using 7-zip compression with an observed overhead on CPU and a bit of memory. Hence it is labelled as a CPU benchmark or CPU intensive workload. 7zip is designed to compress a large file during which the CPU demand remains constantly high. On the other hand, aio-stress constantly reads and writes a 2048 MB test file and a 64KB block size from and to the hard disk. Hence aio-stress is used to test performance and stability of the hard desk. By using 7zip and aio-stress, we have targeted to measure the performance interference resulting from consolidation CPU and disk intensive workloads.

For infrastructure, we have used a server and VM as described in Table 1. 8 VMs are created in a physical server (host), which also runs a PTS Phoromatic server. Both the host and guest run Ubuntu 18.04 server operating system. PTS Phoromatic server makes it possible and easier to execute workload in VMs simultaneously just by a click of a button. The architecture of executing 7zip and aio-stress in server VM is shown in Figure 1.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

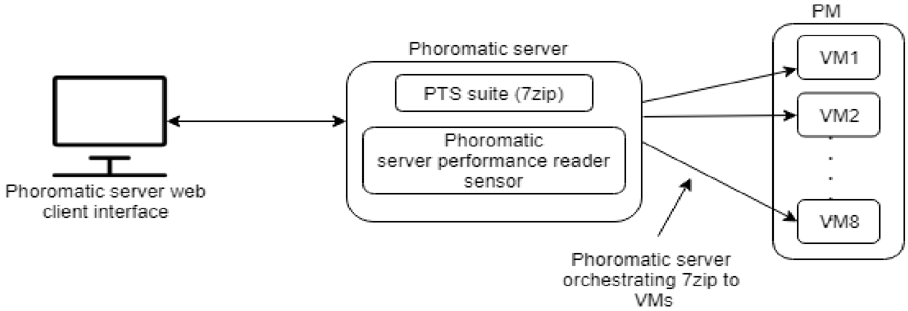

FIGURE 1. Architecture for executing PTS test suits using phoromatic server

TABLE 1.
Characteristics of host and Virtual machines

Item	Item details
Physical server characteristics	Processor: Intel Core i7-4790 @ 4.00GHz, 8 cores Operating system: Ubuntu, Memory size: 12288MB Disk size: 1000GB
Virtual machine characteristics	Processor: 1 vCPU @ 3.59GHz Operating system: Ubuntu 18.04, 4.15.0-36-generic (x86_64) kernel version Memory size: 1280MB Disk size: 45 GB virtual layer: KVM
Number virtual machines	8

7zip and aio-stress is first executed in only one VM in turn and the performance of the CPU and disk I/O measured. The performance of CPU is measured in Million Instructions per Second (MIPS) and that of disk I/O in Megabits per Second (MB/s). Running 7zip and aio-stress in only one VM in tunr gives the base performance of CPU and disk I/O, i.e. when there is no competition from other co-resident VMs. The number VMs running 7zip increases by 1 up to 8 and each time, the average CPU performance is computed. A similar procedure is repeated for aio-stress. The number of VMs ran at every execution for both 7zip and aio-stress represents level consolidation. At each level of consolidation, performance interference index, *PII*, for both the CPU and disk I/O is computed a show in Equation 1.

$$\boldsymbol{PII} = \sum_{i=1}^{k} \frac{P-B}{P} \quad (1)$$

where k is the to total number of VMs running in the physical server, P is the base performance of VM and B is the performance of the VM when it runs with co-resident VMs. A plot of PI against the level of consolidation i.e. number of VMs executing 7zip or aio-stress is plotted in the same chart.

4. RESULTS AND DISCUSSION

Table 2 summarizes CPU and disk I/O performance on 7zip and aio-stress respectively at different levels of performance. The plot shown in Figure. 2 is drawn from data in Table 2 and it shows the effect of increasing the number of co-resident VMs processing workload in the physical server.

TABLE 2.
Performance of CPU at different levels of consolidation

No. of VMs running 7zip	Average CPU performance (MIPS)	7zip PII	Average disk I/O performance (MB/s)	aio-stress PI
1	3737	0	57.57	0
2	3329	0.22	15.38	1.46
3	3215.67	0.42	10.57	2.45
4	2985	0.8	10.35	3.28
5	2890	1.14	8.99	4.29
6	2606.17	1.81	6.55	5.32
7	2518.57	2.28	3.9	6.51
8	2071.25	3.09	3.98	7.71

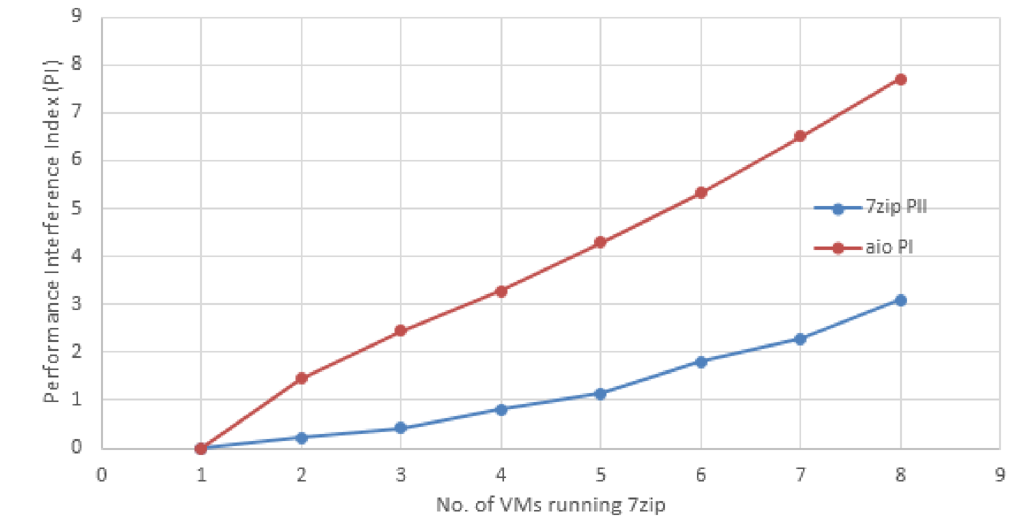

FIGURE 2. The effect of level of consolidation on performance interference

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

The results obtained shows that performance is highest when a single VM is executed in a physical machine for both CPU intensive workloads and disk I/O workloads. This is so because there is no competition for computing resources available in the physical layer. Additionally, there is plenty of resources hence no resource contention. As the number of VMs continue to increase, the average VM performance decreases. This is caused by competition of the hypervisor capacity (in the virtual layer), which makes it possible for the VMs to share physical resources. Increased competition on resources increases performance interference as shown in Figure 2. In fact, performance interference is more severe in homogeneous workloads, which is the case in our experiment because all VM running 7zip were mostly using one resource, CPU and when running aio-stress, the VMs were mostly using disk I/O. Homogenous workloads behave so because they create a hotspot of activity while the rest of the resources remain idle. Futher, when performance interference resulting from consolidating CPU intensive workloads is compared to that in disk I/O intensive workloads, it is observed that the interference is severe in disk I/O intensive workloads.

5. CONCLUSION

In this paper, we have proposed a simple but effective way for measuring inter-VM workload interference applicable in multi-tenant IaaS public cloud. To this end, we have shown that aggressive consolidation of workloads is likely to increase performance interference in co-located VMs and is severe in homogenous workloads. The severity also depends on the type of computing resource that is used. As future work, we plan to compare performance interference when consolidating workloads, which utilize other computing resources such as memory and network I/O. Further, we plan to compare performance interference in homogenous workloads and heterogeneous.

REFERENCES

[1] S. Amri, H. Hamdi and Z. Brahmi, "Inter-VM Interference in Cloud Environments: A Survey," in *2017 IEEE/ACS 14th International Conference on Computer Systems and Applications*, Hammamet, Tunisia, 2017.

[2] X. Pu, L. Liu, Y. Mei, S. Sivathanu, Y. Koh and C. Pu, "Understanding Performance Interference of I/O Workload in Virtualized Cloud Environments," in *2010 IEEE 3rd International Conference on Cloud Computing*, Miami, FL, USA, 2010.

[3] R. Nathuji, A. Kansal and A. Ghaffarkhah, "Q-clouds: managing performance interference effects for QoS-aware clouds," in *EuroSys '10 Proceedings of the 5th European conference on Computer systems*, Paris, France, 2010.

[4] F. Xu, F. Liu and H. Jin, "Heterogeneity and Interference-Aware Virtual Machine Provisioning for Predictable Performance in the Cloud," *IEEE Transactions on Computers*, vol. 65, no. 8, pp. 2470 - 2483, 2016.

[5] X. Chen, L. Rupprecht, R. Osman, P. Pietzuch, F. Franciosi and W. Knottenbelt,

"CloudScope: Diagnosing and Managing Performance Interference in Multi-tenant Clouds," in *2015 IEEE 23rd International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems*, 2015.

[6] Y. Amannejad, D. Krishnamurthy and B. Far, "Detecting performance interference in cloud-based web services," in *2015 IFIP/IEEE International Symposium on Integrated Network Management (IM)*, Ottawa, ON, Canada, 2015.

[7] Phoronix Test Suite , "Phoronix Test Suite Suites," Phoronix Media, 2018. [Online]. Available: https://openbenchmarking.org/suites/pts. [Accessed 05 August 2018].

[8] Phoronix Test Suite, "Phoromatic: Automated Linux Benchmark Management & Test Orchestration," 2018. [Online]. Available: http://www.phoronix-test-suite.com/index.php?k=phoromatic#phoromatic.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)